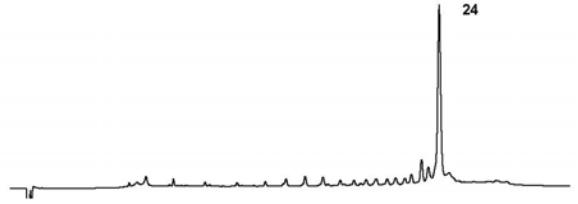
Supporting Information to **'The chemical synthesis of selenium-modified oligoribonucleotides and their enzymatic ligation leading to an U6 snRNA stem-loop segment'**

*Claudia Höbartner and Ronald Micura**

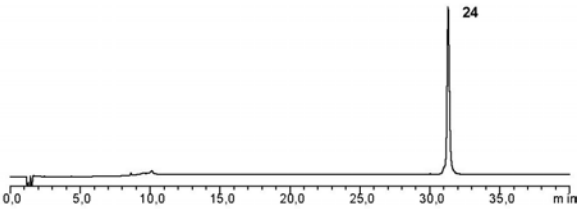
Oligoribonucleotide deprotection.

5'-AAGC$_{Se}$CACACAAACC(dA)(dG)(dA)CGGCC, 24

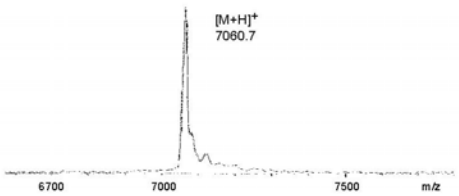
deprotection:
1) $MeNH_2$ 33% in EtOH / $MeNH_2$ 41% in H_2O,
+ **50 mM dithiothreitol**, rt, 6h
2) 1M TBAF in THF, rt, 12h



purified oligoribonucleotide **24**

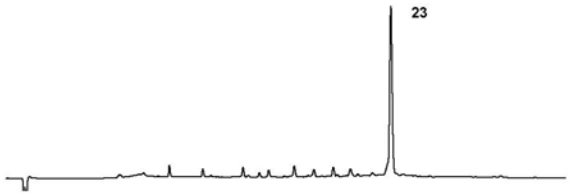


MALDI TOF mass spectrum of **24**

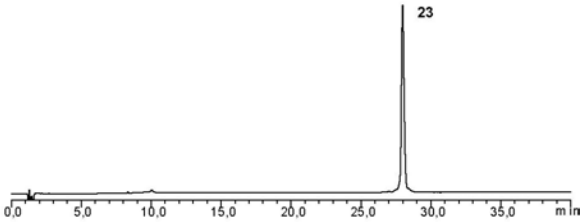


5'-Anthracene-(HEG)-GGAGCU$_{Se}$CGCC$_{Se}$C, 23

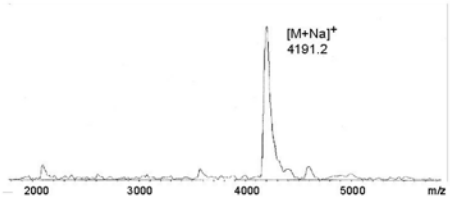
deprotection:
1) $MeNH_2$ 33% in EtOH / $MeNH_2$ 41% in H_2O,
+ **150 mM dithiothreitol**, rt, 5h
2) 1M TBAF in THF, rt, 12h



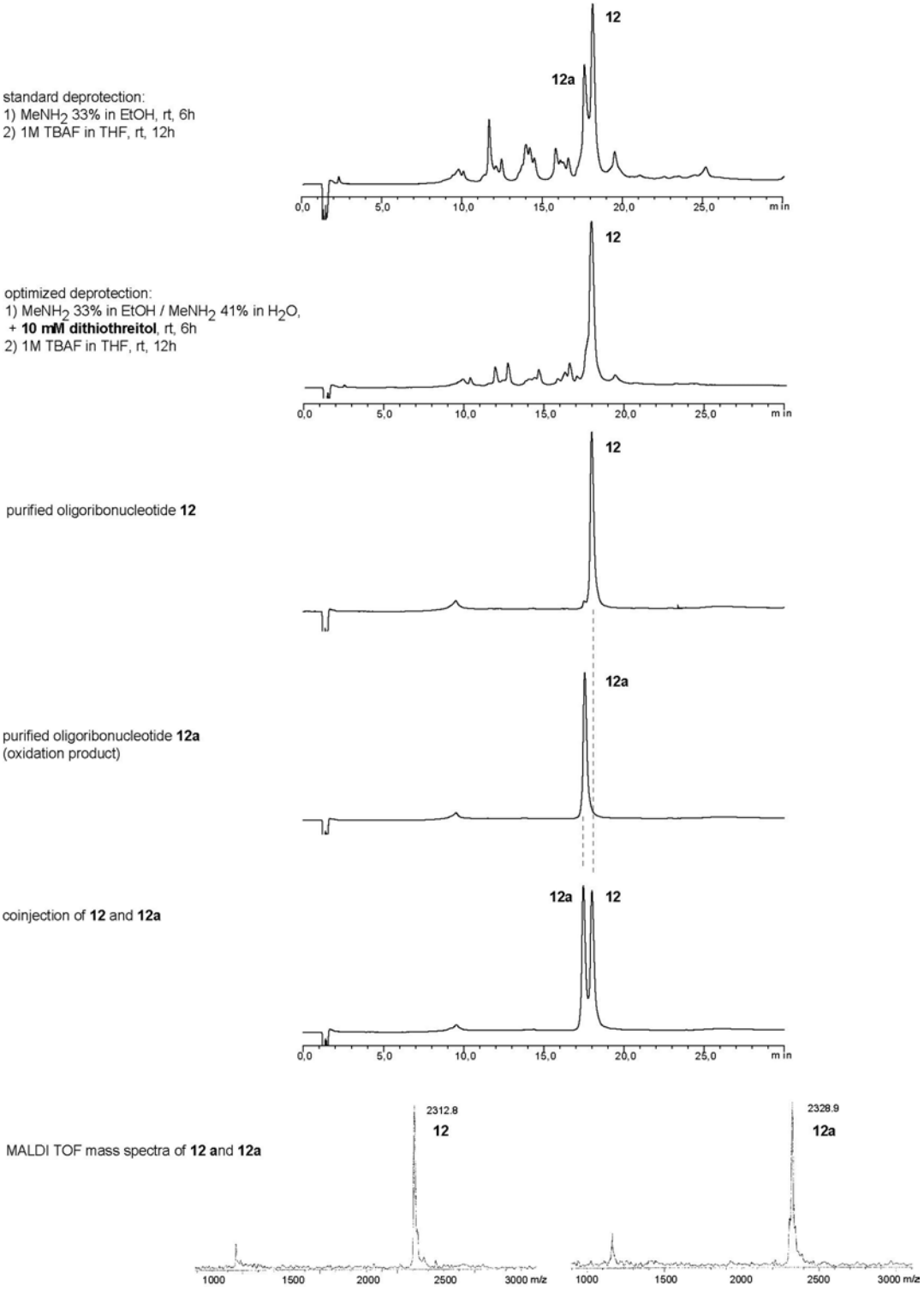
purified oligoribonucleotide **23**



MALDI TOF mass spectrum of **23**



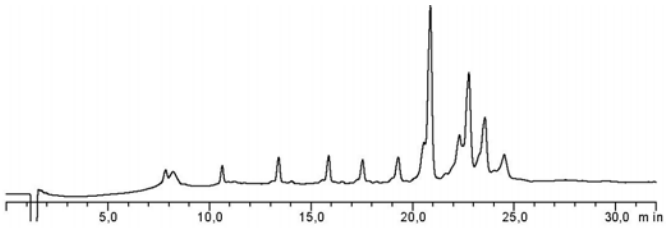
5'-CGCGU$_{Se}$GG, 12



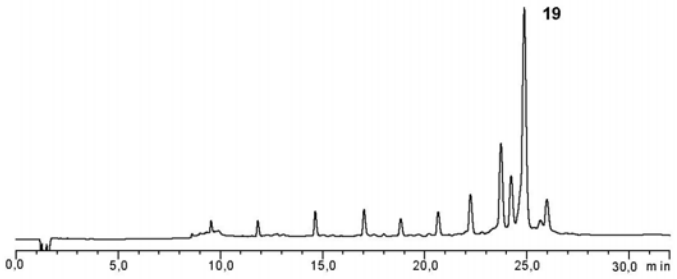
(see also Figure 1)

5'-G C_{Se}GG CGG CGG C, 19

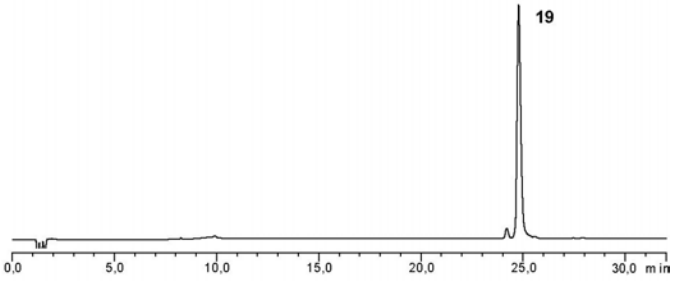
deprotection:
1) NH_3 in MeOH, 42°C, 20h
2) 1M TBAF in THF, rt, 12h



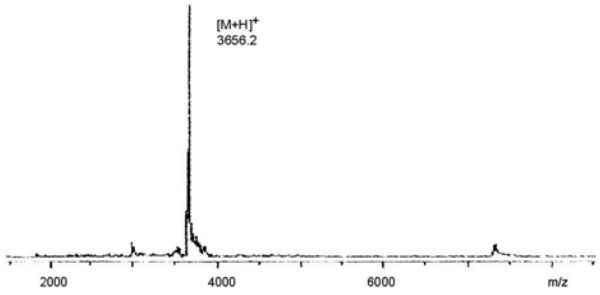
deprotection:
1) $MeNH_2$ 33% in EtOH / $MeNH_2$ 41% in H_2O,
+ **100 mM dithiothreitol**, rt, 8h
2) 1M TBAF in THF, rt, 12h



purified oligoribonucleotide **19**

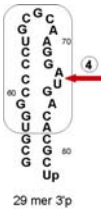


MALDI TOF mass spectrum of **19**

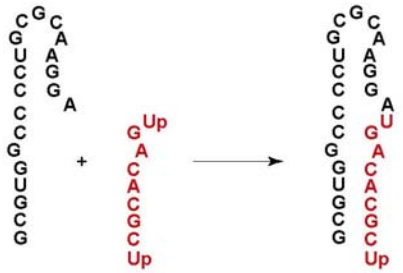


Enzymatic ligations using T4 RNA ligase.

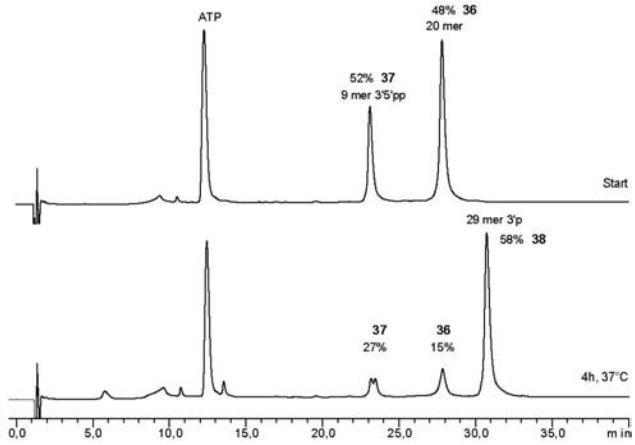
Ligation site 4



unmodified fragment oligoribonucleotides

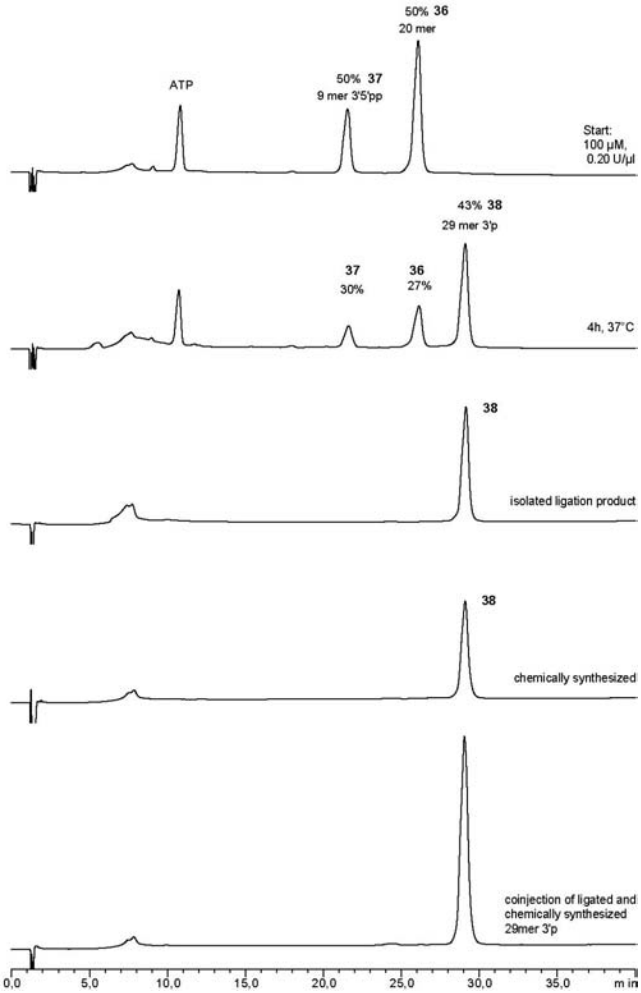


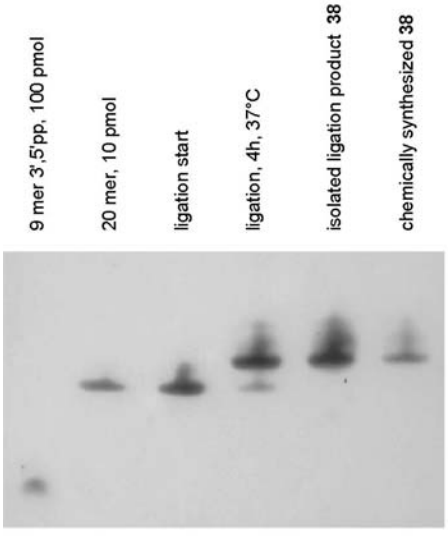
ligation conditions:
2 nmol per oligonucleotide, 40 µM in 1x ligation buffer,
10 U T4 RNA ligase, incubation at 37°C, 4h



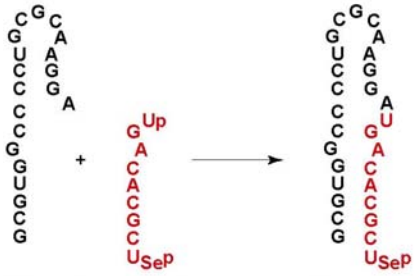
preparative ligation conditions:
5-10 nmol per oligonucleotide, 92-106 µM in
1x ligation buffer, 10 U T4 RNA ligase,
incubation at 37°C, 4h;

calculated yield 43-56 %,
isolated yield: 37 %
(7.5 nmol starting from 20 nmol each)





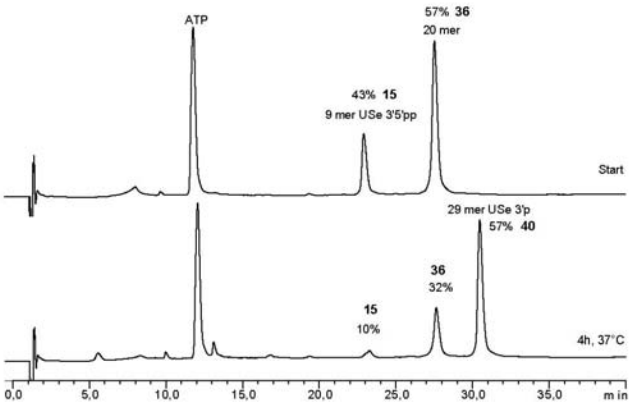
3'-U_{Se}-modified donor oligoribonucleotide



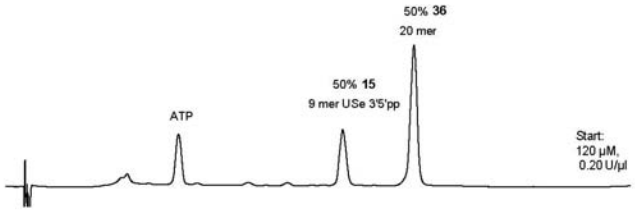
20 mer 9 mer USe 3',5'pp 29 mer 3'p

36 **15** **40**

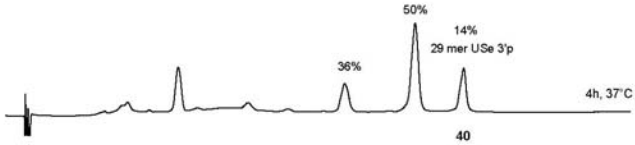
ligation conditions:
2 nmol per oligonucleotide, 40 µM in 1x ligation buffer,
10 U T4 RNA ligase, incubation at 37°C, 4h



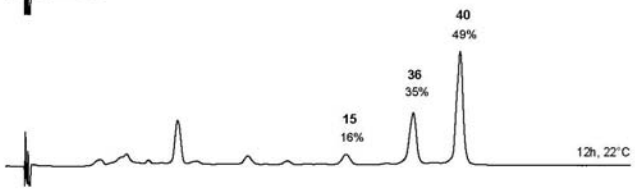
preparative ligation conditions:
10 nmol per oligonucleotide, 120 µM in
1x ligation buffer



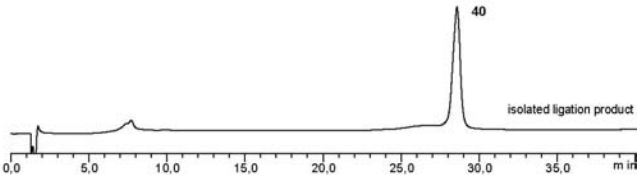
0.2 U/µl T4 RNA ligase,
incubation at 37°C, 4h;
calculated 14% ligation



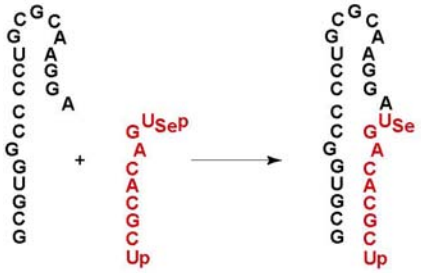
fianlly 0.3 U/µl T4 RNA ligase,
additional incubation at 22°C, 12h;
calculated 49% ligation



calculated yield 49-56%,
isolated yield: 44%
(8.8 nmol starting from 20 nmol each)



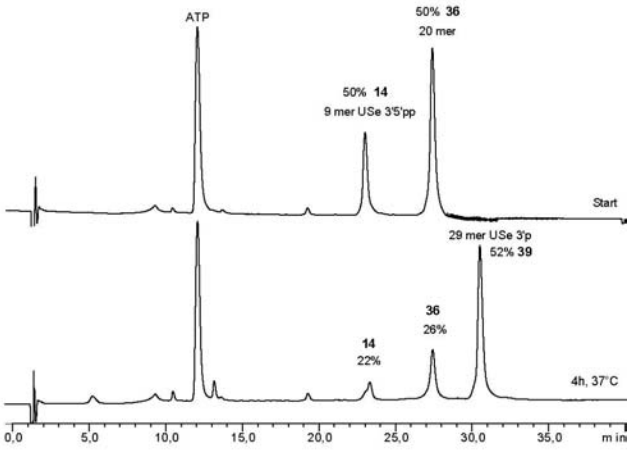
5'-U_{Se}-modified donor oligoribonucleotide



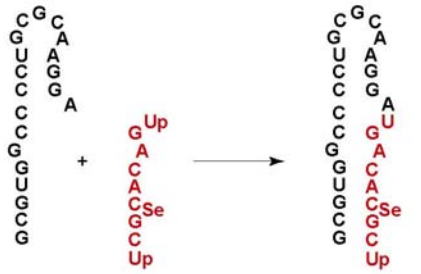
20 mer 9 mer USe 3',5'pp 29 mer 3'p

36 **14** **39**

ligation conditions:
2 nmol per oligonucleotide, 40 µM in 1x ligation buffer,
10 U T4 RNA ligase, incubation at 37°C, 4h



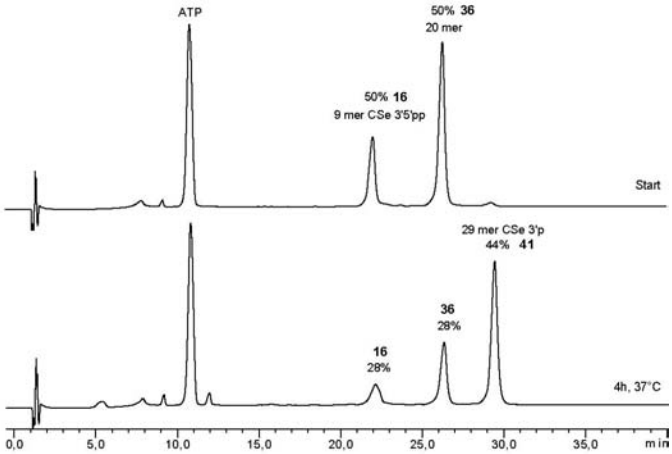
C_{Se}-modified donor oligoribonucleotide



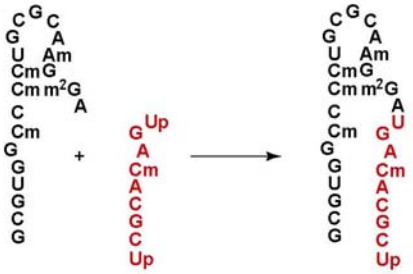
20 mer 9 mer CSe 3',5'pp 29 mer 3'p

36 **16** **41**

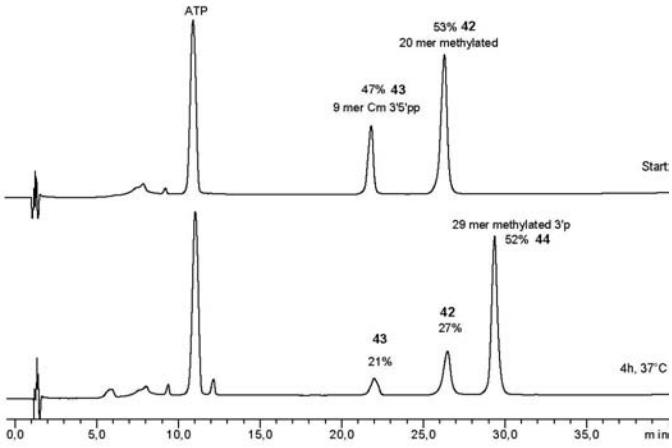
ligation conditions:
2 nmol per oligonucleotide, 40 µM in 1x ligation buffer,
10 U T4 RNA ligase, incubation at 37°C, 4h



methylated fragment oligoribonucleotides

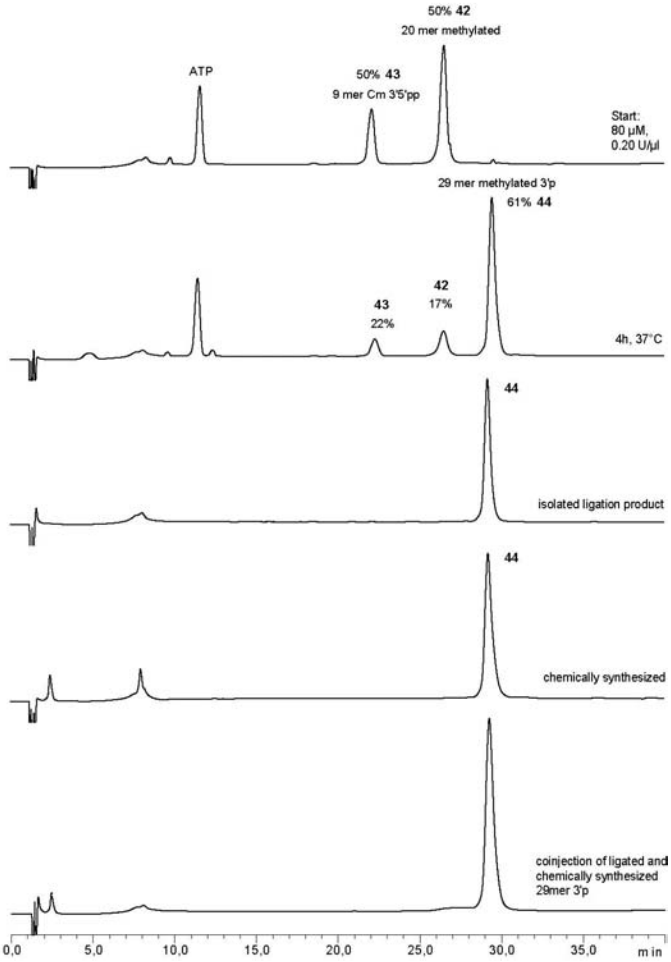


ligation conditions:
2 nmol per oligonucleotide, 40 µM in 1x ligation buffer,
10 U T4 RNA ligase, incubation at 37°C, 4h

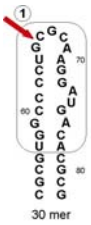


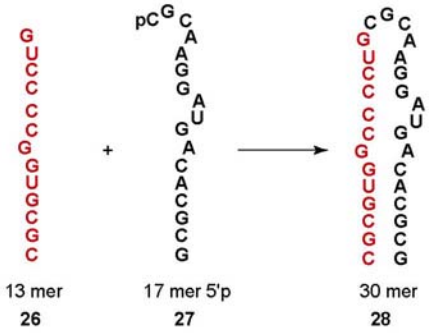
preparative ligation conditions:
5-8 nmol per oligonucleotide, 80 µM in
1x ligation buffer, 10 U T4 RNA ligase,
incubation at 37°C, 4h;

calculated yield 39-61%,
isolated yield: 35%
(3.5 nmol starting from 10 nmol each)

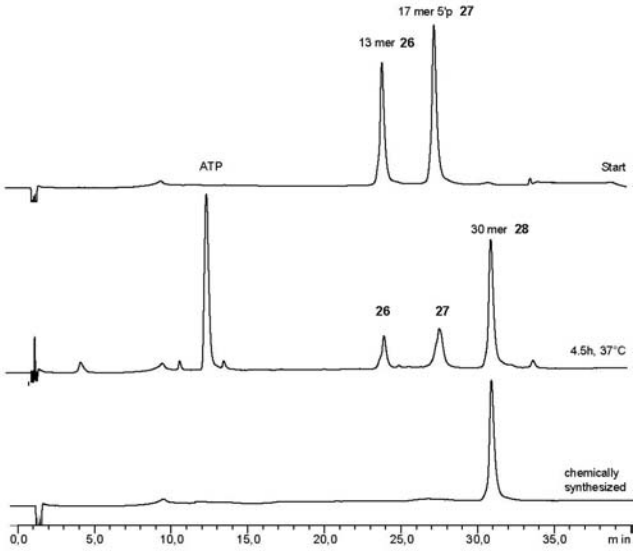


Ligation site 1





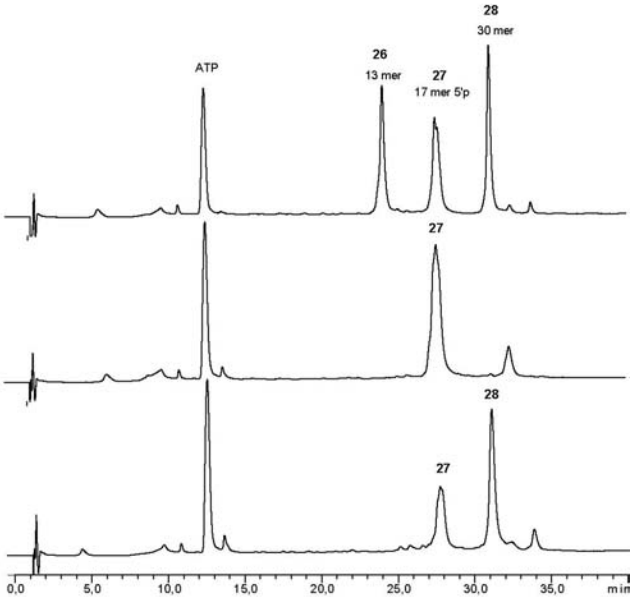
ligation conditions:
2 nmol per oligonucleotide, 40 µM in 1x ligation buffer,
10 U T4 RNA ligase, incubation at 37°C, 4.5h



5 nmol per oligonucleotide,
100 µM in 1x ligation buffer,
10 U T4 RNA ligase, incubation at 37°C, 3.5h

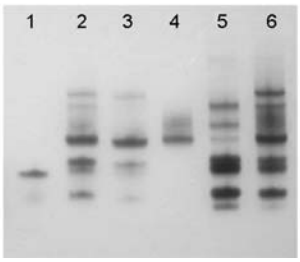
solely 4 nmol donor oligonucleotide **27**,
80 µM in 1x ligation buffer,
10 U T4 RNA ligase, incubation at 37°C, 4.5h

solely 2 nmol per oligonucleotide **27** and **28** (no **26**)
40 µM in 1x ligation buffer,
10 U T4 RNA ligase, incubation at 37°C, 4.5h

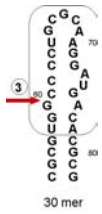


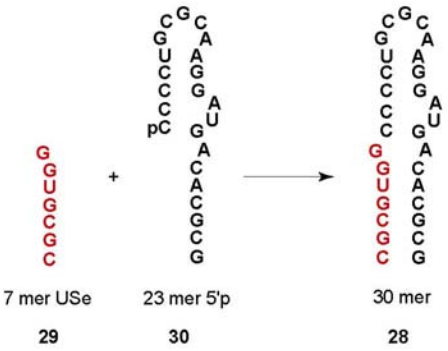
Denaturating Polyacrylamid Gel Electrophoresis

1 equimolar mixture of 13mer (**26**) and 17mer 5'p (**27**), 10 pmol each
2 Ligation **26** + **27**; 3.5h, 37°C; 25 pmol
3 Ligation **26**+ **27**; 3.5h, 37°C; 10 pmol
4 30mer (**28**) from synthesis on solid support
5 Ligation **27** + **27**; 3.5h, 37°C; 20 pmol
6 Ligation **27** + **28**; 3.5h, 37°C; 20 pmol

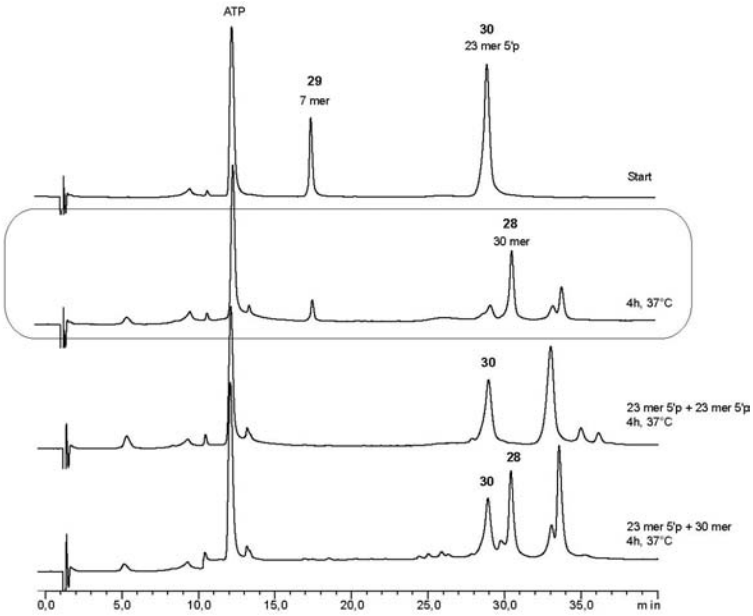


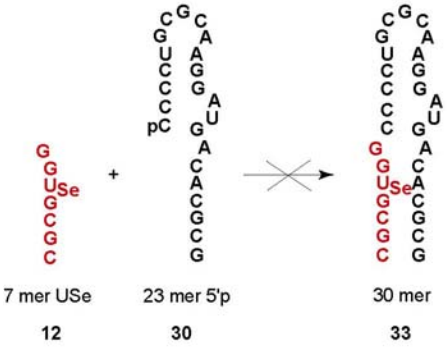
Ligation site 3



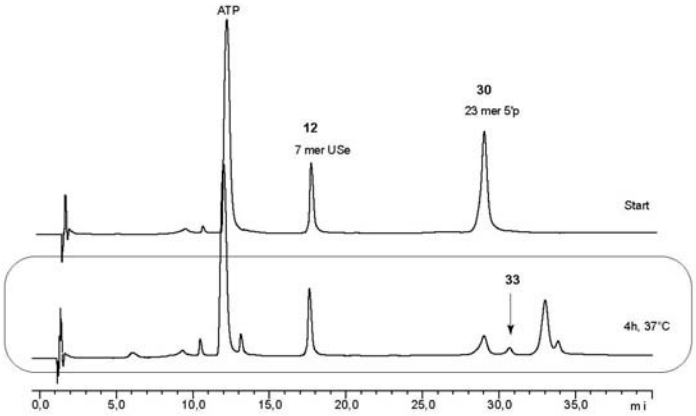


ligation conditions:
2 nmol per oligonucleotide, 40 µM in 1x ligation buffer,
10 U T4 RNA ligase, incubation at 37°C, 4h





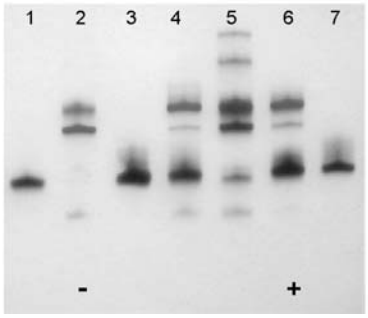
ligation conditions:
2 nmol per oligonucleotide, 40 µM in 1x ligation buffer,
10 U T4 RNA ligase, incubation at 37°C, 4h



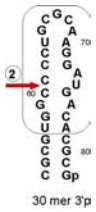
Denaturating Polyacrylamid Gel Electrophoresis

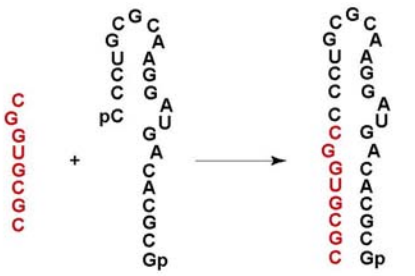
1 equimolar mixture of 7 mer USe (**12**) and 23 mer 5'p (**30**)
2 Ligation 7 mer USe (**12**) + 23 mer 5'p (**30**); 4h, 37°C;
3 equimolar mixture of 23 mer 5'p **30** and 30 mer **28**
4 Ligation 23mer (**30**) + 30 mer (**28**); 4h, 37°C;
5 Ligation 23 mer 5'p (**30**); + 23 mer 5'p (**30**); 4h, 37°C;
6 Ligation 7mer **29** + 23 mer 5'p **30**; 4h, 37°C;
7 chemically synthesized 30mer **28**

line 2: - ... no ligation to **33** observed; only longer byproducts formed
line 6: + ... ligation yields 30mer **28** plus byproducts

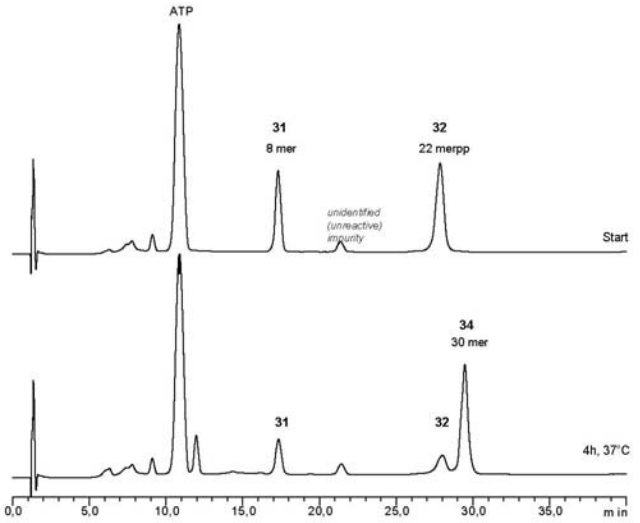


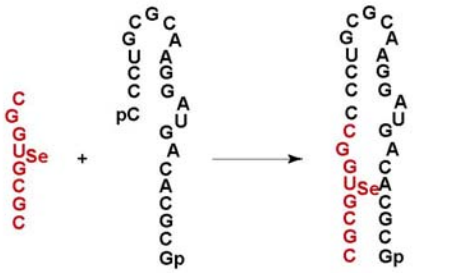
Ligation site 2



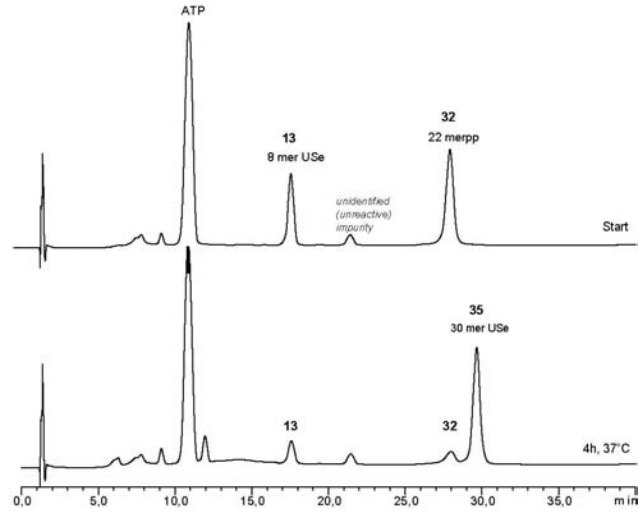


ligation conditions:
1 nmol per oligonucleotide, 20 µM in 1x ligation buffer,
10 U T4 RNA ligase, incubation at 37°C, 4h



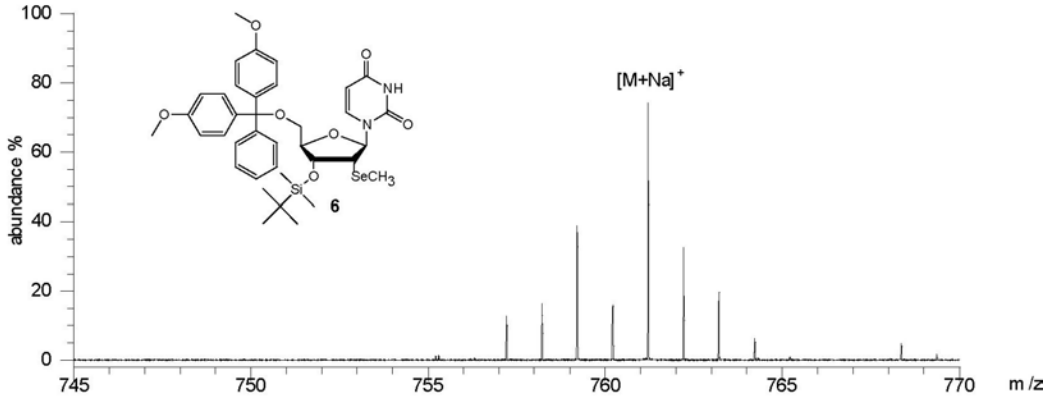


ligation conditions:
1 nmol per oligonucleotide, 20 µM in 1x ligation buffer,
10 U T4 RNA ligase, incubation at 37°C, 4h

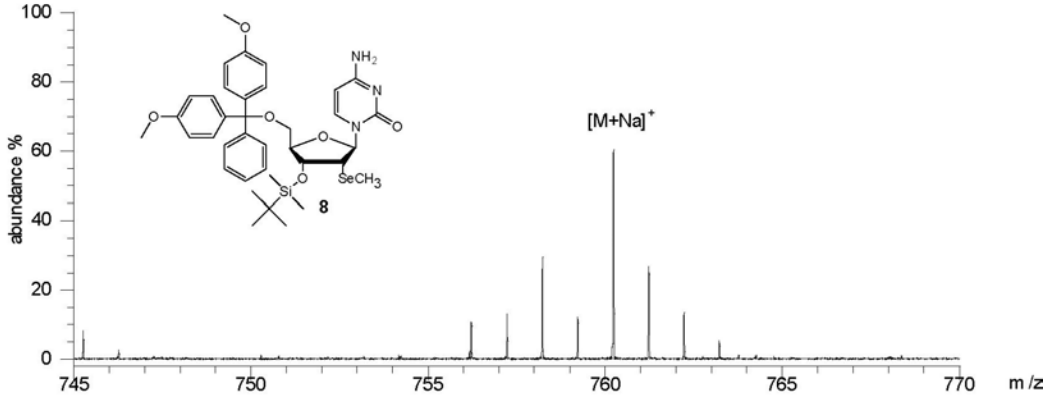


FT-ICR ESI MS spectra of compounds **6**, **8** – **11**.

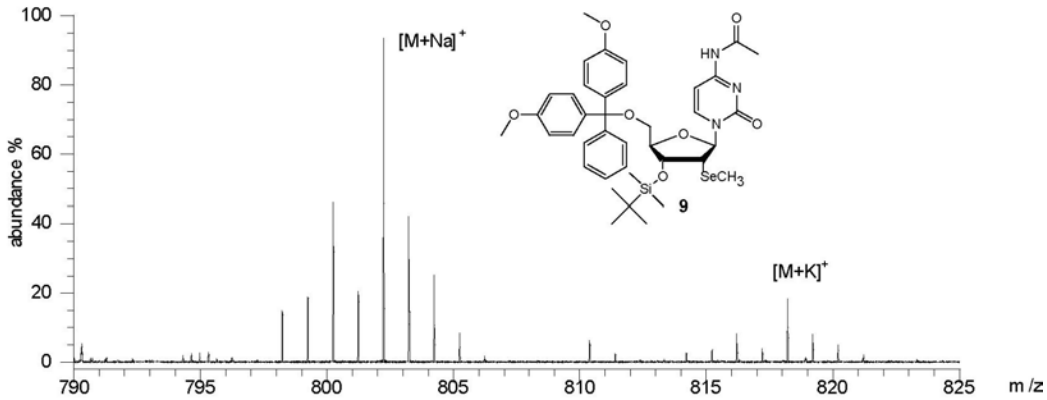
3'-*O*-*Tert*.-butyldimethylsilyl-5'-*O*-(4,4'-dimethoxytrityl)- 2'-*O*-deoxy-2'-*Se*-methyl-uridine (6) FT ICR ESI-MS m/z calculated for $C_{37}H_{46}N_2O_7SeSi$ $[M+ Na]^+$ 761.21421, found 761.21496.



3'-*O*-*Tert*.-butyldimethylsilyl-5'-*O*-(4,4'-dimethoxytrityl)- 2'-*O*-deoxy-2'-*Se*-methyl-cytidine (8) FT ICR ESI-MS m/z calculated for $C_{37}H_{47}N_3O_6SeSi$ $[M+ Na]^+$ 760.23018, found 760.22880.

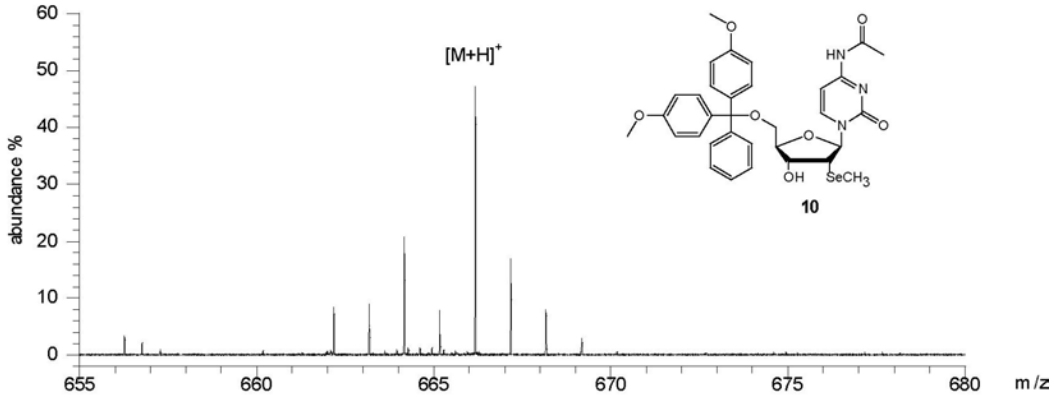


***N*^{4}-Acetyl-3'-*O*-*Tert*.-butyldimethylsilyl-5'-*O*-(4,4'-dimethoxytrityl)- 2'-*O*-deoxy-2'-*Se*-methyl-cytidine (9)** FT ICR ESI-MS m/z calculated for $C_{39}H_{49}N_3O_7SeSi$ $[M+ Na]^+$ 802.24080, found 802.24121.



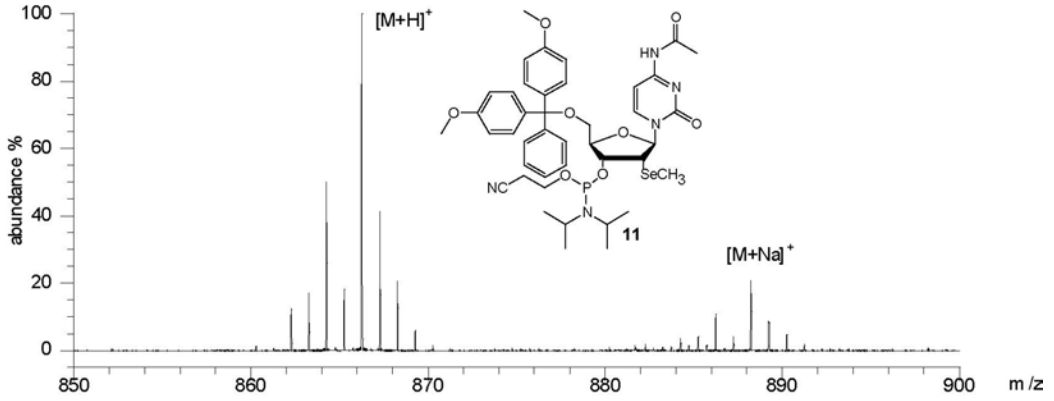
N^4-Acetyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-*O*-deoxy-2'-*Se*-methyl-cytidine (10)

FT ICR ESI-MS m/z calculated for $C_{33}H_{35}N_3O_7SeSi$ $[M+H]^+$ 666.17217, found 666.17208.



N^4-Acetyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-*O*-deoxy-2'-*Se*-methyl-cytidine 3'-O-(2-Cyanoethyl)-diisopropylphosphoramidite (11)

FT ICR ESI-MS m/z calculated for $C_{42}H_{52}N_5O_8PSe$ $[M+H]^+$ 866.28022, found 866.27866.



Representative trityl assays for strand assembly

Synthesis protocol for sequence **5'-CGCGU$_{Se}$GGC, 13**

Gene Assembler Plus

Date : 11/03/2003
Sequence : CH0344-1
Synthesis : CH0344-1
Sequence Length : 8
Column : 1
Final Detritylation : Yes
Coupling Efficiency Threshold : 75 %

Pos	Base	Retention mins	Duration mins	Peak ht %FS	Acc Area %min	Last eff %	Ave eff %
8	C	0.44	1.18	1645	215.89	-	-
7	G	0.36	0.98	1784	212.86	-	-
6	G	0.37	0.95	1778	211.05	99.1	99.1
5	U_{Se}	0.40	1.21	1781	242.21	-	-
4	G	0.37	0.94	1767	209.10	99.5	99.3
3	C	0.43	1.27	1716	251.81	-	99.3
2	G	0.37	1.00	1771	205.21	99.1	99.3
1	C	0.43	1.19	1712	240.58	97.7	98.9

Total synthesis yield from start = 92.4 %

Synthesis protocol for sequence **5'-AAGC$_{Se}$CACACAAACC(dA)(dG)(dA)C$_{Se}$GGCC, 25**

Gene Assembler Plus

Date : 14/07/2003
Sequence : CH0380-4
Synthesis : CH0380-4
Sequence Length : 22
Column : 1
Final Detritylation : Yes
Coupling Efficiency Threshold : 75 %

Pos	Base	Retention mins	Duration mins	Peak ht %FS	Acc Area %min	Last eff %	Ave eff %
22	C	0.43	1.42	1781	279.10	-	-
21	C	0.44	1.35	1761	312.34	-	-
20	G	0.37	0.89	1767	218.05	-	-
19	G	0.38	0.98	1775	231.08	106.0	100.0
18	C_{Se}	0.43	1.20	1764	267.29	-	100.0
17	dA	0.42	1.26	1767	250.28	-	-
16	dG	0.42	1.02	1758	232.96	-	-
15	dA	0.41	1.18	1764	239.14	-	-
14	C	0.45	1.24	1708	259.70	97.4	100.0
13	C	0.47	1.30	1650	263.98	101.6	100.0
12	A	0.45	1.17	1690	239.73	-	100.0
11	A	0.46	1.20	1660	240.83	100.5	100.0
10	A	0.44	1.14	1691	229.59	95.3	100.0
9	C	0.46	1.29	1611	239.61	97.6	99.7
8	A	0.44	1.12	1660	221.05	98.1	99.5
7	C	0.45	1.26	1586	229.10	97.8	99.3
6	A	0.44	1.18	1634	215.88	98.8	99.2
5	C	0.46	1.34	1517	228.10	99.8	99.3
4	C_{Se}	0.45	1.27	1580	215.01	98.5	99.2
3	G	0.40	0.98	1589	176.14	98.3	99.1
2	A	0.44	1.23	1488	193.60	97.3	99.0
1	A	0.45	1.07	1462	189.16	97.7	98.9

Total synthesis yield from start = 79.4 %